On the Distributions of the Lengths of the Longest Monotone Subsequences in Random Words

Craig A. Tracy
Department of Mathematics and Institute of Theoretical Dynamics
University of California, Davis, CA 95616, USA
e-mail address: tracy@itd.ucdavis.edu

Harold Widom
Department of Mathematics
University of California, Santa Cruz, CA 95064, USA
e-mail address: widom@math.ucsc.edu

Abstract

We consider the distributions of the lengths of the longest weakly increasing and strongly decreasing subsequences in words of length N from an alphabet of k letters. (In the limit as $k \to \infty$ these become the corresponding distributions for permutations on N letters.) We find Toeplitz determinant representations for the exponential generating functions (on N) of these distribution functions and show that they are expressible in terms of solutions of Painlevé V equations. We show further that in the weakly increasing case the generating function gives the distribution of the smallest eigenvalue in the $k \times k$ Laguerre random matrix ensemble and that the distribution itself has, after centering and normalizing, an $N \to \infty$ limit which is equal to the distribution function for the largest eigenvalue in the Gaussian Unitary Ensemble of $k \times k$ hermitian matrices of trace zero.

I. Introduction

The last decade has seen a flurry of activity centering around connections between combinatorial probability on the one hand and random matrices and integrable systems on the other. From the point of view of probability theory, the quite surprising feature of these developments is that the methods came from Toeplitz determinants, integrable differential equations of the Painlevé type and the closely related Riemann-Hilbert techniques as they were applied and refined in random matrix theory. Using these methods new, and apparently quite universal, limiting laws have been discovered. One of the aims of this paper is to make these methods accessible to a wider audience. Our story begins with a theorem of Gessel [16]. (There are earlier signs of these connections; see Regev [30].)

Let S_N be the symmetric group on N letters and give each permutation $\sigma \in S_N$ probability $1/N!$. Denote by $\ell_N(\sigma)$ the length of the longest increasing subsequence in σ and

$$F_P(n; N) = \mathrm{Prob}\left(\ell_N(\sigma) \le n\right).$$

Then it is a corollary of Gessel's theorem and the Robinson-Schensted-Knuth correspondence[1] that

$$\sum_{N=0}^{\infty} F_P(n;N)\frac{t^N}{N!} = D_n(t), \tag{1.1}$$

where $D_n(t)$ is the determinant of the the $n \times n$ Toeplitz matrix with the symbol $e^{\sqrt{t}\,(z+z^{-1})}$. (Recall that the i, j entry of a Toeplitz matrix equals the $i-j$ Fourier coefficient of its symbol.)

It is in this work of Gessel, expressing the (exponential) generating function of F_P as a Toeplitz determinant, and the subsequent work of Odlyzko *et al.* [26] and Rains [29], that the methods of random matrix theory first appear in RSK type problems.[2]

Starting with this representation, Baik, Deift and Johansson [2], using the steepest descent method for Riemann-Hilbert problems [12], derived a delicate asymptotic formula for $D_n(t)$ which we now describe. Introduce another parameter s and suppose that n and t are related by $n = [2t^{1/2} + st^{1/6}]$. Then as $t \to \infty$ with s fixed one has

$$\lim_{t\to\infty} e^{-t} D_{2\,t^{1/2}+s\,t^{1/6}}(t) = F_2(s).$$

Here F_2 is the distribution function defined by

$$F_2(s) = \exp\left(-\int_s^{\infty}(x-s)q(x)^2\,dx\right) \tag{1.2}$$

where q is the solution of the Painlevé II equation

$$q'' = sq + 2q^3$$

satisfying $q(s) \sim \mathrm{Ai}(s)$ as $s \to \infty$.[3] Using a dePoissonization lemma due to Johansson [20], these asymptotics for $D_n(t)$ led Baik, Deift and Johansson to the limiting law

$$\lim_{N\to\infty} \mathrm{Prob}\left(\frac{\ell_N(\sigma) - 2\sqrt{N}}{N^{1/6}} \le s\right) = F_2(s).$$

It is a remarkable fact that this same distribution function F_2 was first encountered by the present authors [34] in random matrix theory where it arises as the limiting law for the normalized largest eigenvalue in the Gaussian Unitary Ensemble (GUE) of Hermitian matrices. More precisely, we have for this ensemble [34]

$$\lim_{N\to\infty} \mathrm{Prob}\left((\lambda_{\max} - \sqrt{2N})\sqrt{2}N^{1/6} \le s\right) = F_2(s). \tag{1.3}$$

Here we see a connection with integrable systems—the appearance of a Painlevé II function. Yet another connection is that $D_n(t)$ itself has a representation in terms of the solution of a Painlevé V equation [18, 37].

[1] This is a bijection between permutations and pairs (P, Q) of standard Young tableaux with the same shape. For expository accounts of the RSK algorithm, see [15, 24, 33].

[2] Gessel [16] does not mention random matrices, but in light of well-known formulas in random matrix theory relating Toeplitz determinants to expectations over the unitary group, we believe it is fair to say that the connection with random matrix theory begins with his discovery.

[3] Ai is the Airy function. For a proof that such a solution exists and is unique, see [8, 17, 13].

Since the work of Baik, Deift and Johansson, several groups have extended this connection between RSK type combinatorics and the distribution functions of random matrix theory. The aforementioned result is equivalent to the determination of the limiting distribution of the number of boxes in the first row in the RSK correspondence $\sigma \leftrightarrow (P, Q)$. In [3] the same authors show that the limiting distribution of the number of boxes in the *second* row is (when centered and normalized) distributed as the *second* largest scaled eigenvalue in GUE [34]. They then conjectured that this correspondence extends to all rows. This conjecture was recently proved by Okounkov [28] using topological methods and by Borodin, Okounkov and Olshanski [7] using analytical/representation theoretic methods.

Placing restrictions on the permutations σ (that they be fixed point free and involutions), Baik and Rains [4] have shown that the limiting laws for the length of the longest increasing/decreasing subsequence are now the limiting distributions F_1 and F_4 [36] for the scaled largest eigenvalues in the Gaussian Orthogonal Ensemble (GOE) and the Gaussian Symplectic Ensemble (GSE). Generalizing to signed permutations and colored permutations the present authors and Borodin [37, 6] showed that the distribution functions of the length of the longest increasing subsequence involve the same F_2. Johansson [21] showed that the shape fluctuations of a certain random growth model, again appropriately scaled, converges in distribution to F_2. (This random growth model is intimately related to certain randomly growing Young diagrams.)

Finally, we mention the work of Aldous and Diaconis [1] where they describe a certain one-person card game, called "patience sorting," which is connected to these ideas linking Young tableaux with the length of the longest increasing subsequence in either a random permutation or a random word; and thence to the limiting distributions of largest eigenvalues.

At last we come to the subject of the present paper, which is the question of what can be said when instead of a random permutation on N letters we have a random word of length N from an alphabet of k letters. This may be thought of as a function from $\{1, 2, \cdots, N\}$ to $\{1, 2, \cdots, k\}$ and it is clear what is meant by a (strictly or weakly) increasing or decreasing subsequence. Unlike the case of permutations there is a difference between the two. Given such a word w we define $\ell_N^I(w)$ to be the length of the longest *weakly increasing* subsequence in w, and define $\ell_N^D(w)$ to be the length of the longest *strictly decreasing* subsequence in w.[4] In analogy with permutations we define the distribution functions (giving each word probability k^{-N})

$$F_I(n; k, N) = \mathrm{Prob}_k\left(\ell_N^I(w) \le n\right), \quad F_D(n; k, N) = \mathrm{Prob}_k\left(\ell_N^D(w) \le n\right)$$

and their generating functions

$$G_I(n; k, t) = \sum_{N=0}^{\infty} F_I(n; k, N)\,\frac{t^N}{N!}, \quad G_D(n; k, t) = \sum_{N=0}^{\infty} F_D(n; k, N)\,\frac{t^N}{N!}. \tag{1.4}$$

Here are our results. We use the standard notation $T_n(f)$ for the $n \times n$ Toeplitz matrix with symbol f.

[4] We could just as well consider weakly increasing subsequences *and* strictly increasing subsequences since reversing the order of the word takes increasing subsequences into decreasing subsequences. In light of the RSK algorithm, our choice seems most convenient.

Theorem 1. We have

$$G_I(n;k,kt) = \det\, T_n(f_I), \quad G_D(n;k,kt) = \det\, T_n(f_D),$$

where

$$f_I(z) = f_I(z;t) = e^{t/z}\,(1+z)^k, \quad f_D(z) = f_D(z;t) = e^{t/z}\,(1-z)^{-k}.$$

Theorem 2. $G_I(n;k,t)$ and $G_D(n;k,t)$ have integral representations in terms of solutions of Painlevé V equations.

Theorem 3. $G_I(n;k,t)$ is equal to e^{kt} times the distribution function for the smallest eigenvalue in the Laguerre ensemble of $k \times k$ matrices associated with the weight function $x^n\, e^{-x}$.

Theorem 4. The limiting distribution for the random variable $\ell_N^I(w)$, centered and normalized, is equal to that for the largest eigenvalue in the Gaussian Unitary Ensemble ensemble of $k \times k$ hermitian matrices with trace zero.[5] More precisely,

$$\lim_{N\to\infty} \mathrm{Prob}_k \left(\frac{\ell_N^I(w) - N/k}{\sqrt{2N/k}} \le s \right) = \mathrm{Prob}\left(\lambda_{\max} \le s\right).$$

The next four sections contain the proofs of these four theorems.

Theorem 1 is a consequence of Gessel's theorem, just as the permutation analogue (1.1) is, and the RSK correspondence between words and pairs of Young tableaux. For the convenience of the reader we include a complete proof of Gessel's theorem, containing the main ideas of the original but presented somewhat differently. (For related developments, see [5].)

Theorem 2 is the heart of the paper. The equations are derived very much in the spirit of [37]. The logarithmic derivative of the determinant involves a quantity whose derivatives in turn involve other quantities. Recursion formulas relating the various quantities allow the eventual derivation of a single differential equation which, in the end, turns out to be reducible to Painlevé V.

The proof of Theorem 3 consists of showing that the P_V function of Theorem 2 for G_I is exactly the one which gives the distribution of the smallest eigenvalue in the Laguerre ensemble [35]. The equation is the same, by inspection, and it is a matter of checking the boundary condition at $t = 0$.

Given the results of [3, 28, 7] for permutations, it is natural to conjecture that the limiting distribution of the number of boxes in the j^{th} row, $2 \le j \le k$, appearing in the Young tableaux P in the RSK bijection $w \leftrightarrow (P,Q)$ is precisely the distribution of the j^{th} largest eigenvalue in the finite $k \times k$ Hermite ensemble.

Theorem 4 is proved by an asymptotic evaluation of the multiple integral which gives the distribution function for the smallest eigenvalue in the Laguerre ensemble.

[5] The normalization we adopt for the GUE measure is the standard one defined in Mehta [25]. The probability on the right of the displayed formula is the conditional probability given that the matrix from GUE has trace zero.

After the completion of the original version of this paper there were several relevant developments. Johansson [22] found an independent proof of Theorem 4, in fact of the full conjecture stated above. A. Its, using Riemann-Hilbert techniques applied to operator equivalents of our Toeplitz matrices, found another proof of Theorem 2. J. L. Snell found an error in the original version of Theorem 4. (We thank him for catching the error and so saving the authors from further embarrassment.) C. Grinstead found a random walk interpretation of the $k = 2$ problem and used this to determine the limiting distribution in this case. And Y. Chen alerted us to the paper [14] of Forrester in which there appeared a formula equivalent (given Theorem 1) to the statement of Theorem 3, obtained by using identities of Macdonald on hypergeometric functions of several variables.

II. Gessel's theorem and its specializations

1. The Cauchy-Binet formula

We begin by recalling the Cauchy-Binet formula for the determinant of the product of two rectangular matrices A and B of sizes $m \times n$ and $n \times m$, respectively. We assume $n \geq m$.

Let $\mathcal{S}_{mn}$ denote the set of strictly increasing subsequences of length m that can be chosen from $\{1, 2, \ldots, n\}$. For any matrix X of size $n \times m$ and any $S = \{s_1, s_2, \ldots, s_m\} \in \mathcal{S}_{mn}$, denote by $X(S|m)$ the $m \times m$ matrix obtained from X by using all m columns of X and the m rows numbered by S. Similarly, if X is $m \times n$, denote by $X(m|S)$ the $m \times m$ matrix obtained from all m rows of X and the columns of X numbered by S. The Cauchy-Binet formula is

$$\det(AB) = \sum_{S \in \mathcal{S}_{mn}} \det\left(B(S|m)\right) \det\left(A(m|S)\right). \tag{2.1}$$

Here is a proof. Define

$$F(\varepsilon) = \varepsilon^m \det\left(I + \varepsilon^{-1} AB\right).$$

This determinant is $m \times m$ and $\lim_{\varepsilon \to 0} F(\varepsilon) = \det(AB)$. Then also

$$F(\varepsilon) = \varepsilon^m \det\left(I + \varepsilon^{-1} BA\right).$$

In the (Fredholm) expansion of this last determinant, the term with coefficient ε^{-m} is

$$\frac{1}{m!} \sum_{i_1, i_2, \ldots, i_m} \begin{vmatrix} (BA)_{i_1 i_1} & (BA)_{i_1 i_2} & \cdots & (BA)_{i_1 i_m} \\ (BA)_{i_2 i_1} & (BA)_{i_2 i_2} & \cdots & (BA)_{i_2 i_m} \\ \cdots & \cdots & \cdots & \cdots \\ (BA)_{i_m i_1} & (BA)_{i_m i_2} & \cdots & (BA) i_m i_m \end{vmatrix}$$

In this sum we can place the restriction that no two indices are equal since when they are the determinant is zero. The $m!$ different orderings of $\{i_1, \ldots, i_m\}$ give the same determinant so we can drop the $m!$ and sum over all i_α with $i_1 < i_2 < \cdots < i_m$. The terms of order ε^{-j}, $j < m$, in the Fredholm expansion do not contribute to the limit as $\varepsilon \to 0$. The coefficients of the terms ε^{-j}, $j > m$, are zero since the rank of BA is at most m. Finally, each summand factors into the product of the two determinants in the Cauchy-Binet formula.

The formula remains valid if $n = \infty$ if, for example, each row of A and each column of B belongs to the sequence space ℓ^2. For then AB is well-defined, BA is a finite rank operator on ℓ^2 and the preceding goes through without change. The sum on the right side of (2.1) is then the sum over all increasing subsequences S of length m chosen from the positive integers.

2. Gessel's theorem

Let $\mathcal{P}_m$ denote the set of partitions of m, sequences $(\lambda_1, \lambda_2, \ldots, \lambda_n)$ of nonnegative integers such that $\lambda_1 \geq \cdots \geq \lambda_n$, $\lambda_1 + \cdots + \lambda_n = m$, and $\mathcal{P} = \bigcup_{m \geq 0} \mathcal{P}_m$. ($\mathcal{P}_0$ is the empty partition.) We also write $\lambda \vdash m$ when $\lambda \in \mathcal{P}_m$ and denote by $\ell(\lambda)$ the length of the partition, the largest k such that $\lambda_k \neq 0$. Let $\Lambda_{\mathbf{Q}}$ (or Λ for short) denote the algebra of symmetric functions over $\mathbf{Q}$. This is a commutative algebra and the vector space direct sum decomposition into homogeneous symmetric functions gives $\Lambda_{\mathbf{Q}}$ the structure of a graded algebra.

Gessel introduces

$$R_n(x,y) := \sum_{\substack{\lambda \in \mathcal{P} \\ \ell(\lambda) \leq n}} s_\lambda(x) s_\lambda(y) = \sum_{m=0}^{\infty} \sum_{\substack{\lambda \in \mathcal{P}_m \\ \ell(\lambda) \leq n}} s_\lambda(x) s_\lambda(y) \tag{2.2}$$

where $s_\lambda(x)$ is the Schur function. Gessel's theorem says that $R_n(x,y)$ is a Toeplitz determinant

$$R_n(x,y) = \det \left(A_{i-j}\right)_{1 \leq i,j \leq n} \tag{2.3}$$

where

$$A_i = A_i(x,y) = \sum_{\ell=0}^{\infty} h_{\ell+i}(x) h_\ell(y)$$

and h_r is the r^{th} complete symmetric function. (We take $h_r = 0$ for $r < 0$.) Recall that

$$\sum_{r \geq 0} h_r t^r = \prod_{i \geq 1} \left(1 - x_i t\right)^{-1}.$$

Here is Gessel's proof. Although (2.3) is an identity between two formal power series it suffices to prove it when the x_i are real numbers satisfying $|x_i| < 1$. Let $M(x)$ be the $\infty \times n$ matrix

$$(h_{i-j}(x)), \qquad (i \geq 1,\ 1 \leq j \leq n).$$

This will be the matrix B of (2.1) whereas A will be $M^t(y)$. (We interchanged the roles of m and n.) Since the entries of the columns of $M(x)$ and rows of $M^t(y)$ are exponentially small, the formula holds.

For any increasing subsequence S of positive integers of length n, let $M_S(x)$ be the determinant of the $n \times n$ minor of $M(x)$ obtained from the rows indexed by the elements of S. In the notation of (2.1) we have $M_S(x) = \det(B(S|n))$, $M_S(y) = \det(A(n|S))$.

Now let $\lambda = (\lambda_1, \ldots, \lambda_n)$ be a partition with $\ell(\lambda) \leq n$, and let $S = \{\lambda_{n+1-i} + i | 1 \leq i \leq n\}$. Observe that S is an increasing subsequence of postive integers. Then

$$M_S(x) = \det \left(h_{\lambda_{n+1-i}+i-j}(x)\right)_{1 \leq i,j \leq n}.$$

Reversing the order of the rows and columns in this determinant yields

$$M_S(x) = \det\left(h_{\lambda_i+j-i}(x)\right) = s_\lambda(x),$$

where the last equality is the Jacobi-Trudi identity. For such S, summing over all partitions λ with $\ell(\lambda) \le n$ is the same as summing over all increasing subsequences S of length n. Thus

$$R_n(x,y) = \sum_S M_S(x) M_S(y),$$

and by the Cauchy-Binet formula this is equal to $\det(M^t(y)\, M(x))$. Since

$$(M^t(y)\, M(x))_{ij} = \sum_\ell h_{\ell-i}(y)\, h_{\ell-j}(x) = \sum_\ell h_{\ell+i-j}(x)\, h_\ell(y) = A_{i-j},$$

the theorem follows.

For the applications which follow it is important to know the the symbol of the Toeplitz determinant appearing in (2.3). It is

$$\begin{aligned}\varphi(z) = \sum_{i=-\infty}^{\infty} A_i(x,y)\, z^i &= \sum_{i=-\infty}^{\infty} z^i \sum_{\ell=0}^{\infty} h_{\ell+i}(x) h_\ell(y) = \sum_{\ell=0}^{\infty} h_\ell(y) \sum_{i=-\infty}^{\infty} h_i(x) z^{i-\ell} \\ &= \prod_{n=1}^{\infty}\left(1 - y_n\, z^{-1}\right)^{-1} \prod_{n=1}^{\infty} (1 - x_n\, z)^{-1}. \qquad (2.4)\end{aligned}$$

3. Cauchy's identity from Szegö's theorem

A nice application of Gessel's theorem is a derivation of Cauchy's identity in symmetric functions[6] using the strong Szegö's limit theorem for Toeplitz determinants. We have $(\log\varphi)_0 = 0$ and, for $n > 0$,

$$(\log\varphi)_n = \frac{1}{n}\sum_{i\ge 1}(x_i)^n,$$

$$(\log\varphi)_{-n} = \frac{1}{n}\sum_{i\ge 1}(y_i)^n.$$

(The subscripts denote Fourier coefficients, as usual.) Applying Szegö's theorem (we may assume that the x_i and y_i are real numbers with absolute value less than 1) then gives

$$\lim_{n\to\infty} R_n(x,y) = \exp\left\{\sum_{n=1}^{\infty}\frac{1}{n}\sum_{i,j\ge 1}(x_i\, y_j)^n\right\} = \prod_{i,j}\left(1 - x_i y_j\right)^{-1},$$

and hence Cauchy's identity

$$\sum_{\lambda\in\mathcal{P}} s_\lambda(x) s_\lambda(y) = \prod_{i,j}\left(1 - x_i y_j\right)^{-1}.$$

[6] We freely use various results from symmetric functions which can be found, for example, in [33], Chp. 7.

4. Dual version of Gessel's theorem

Since (2.3) is an identity between two elements of the ring generated by the x_i and y_i, any endomorphism of this ring yields another identity. Now the complete symmetric functions h_r are algebraically independent generators of Λ as are the elementary symmetric functions e_r. We consider the endomorphism ω defined by

$$\omega(e_r) = h_r.$$

Then for any partition $\lambda = (\lambda_1, \lambda_2, \ldots)$ we have

$$\omega(e_\lambda) = h_\lambda$$

with the usual notation

$$e_\lambda = \prod_i e_{\lambda_i}, \quad h_\lambda = \prod_i h_{\lambda_i}.$$

The action on the Schur function is given by $\omega(s_\lambda) = s_{\lambda'}$ where λ' is the partition conjugate to λ.

We define

$$\tilde{R}_n(x,y) = \omega_x\, \omega_y\, R_n(x,y) = \sum_{\substack{\lambda\in\mathcal{P}\\ \ell(\lambda)\le n}} s_{\lambda'}(x) s_{\lambda'}(y).$$

Now for any partition λ we have $\ell(\lambda) = \lambda'_1$, the length of the first row of the Young diagram of shape λ', and so

$$\tilde{R}_n(x,y) = \sum_{\substack{\lambda\in\mathcal{P}\\ \lambda'_1\le n}} s_{\lambda'}(x) s_{\lambda'}(y) = \sum_{\substack{\lambda\in\mathcal{P}\\ \lambda_1\le n}} s_{\lambda}(x) s_{\lambda}(y).$$

(Thus, in this sum we restrict the length of the first row rather than the length of the first column.) Applying $\omega_x\, \omega_y$ to (2.3) we obtain the dual version of Gessel's theorem,

$$\sum_{\substack{\lambda\in\mathcal{P}\\ \lambda_1\le n}} s_{\lambda}(x) s_{\lambda}(y) = \det\, \left(\omega_x\, \omega_y\, A_{i-j}(x,y)\right)_{i,\,j=1,\ldots,n}.$$

We record for use below

$$\omega_x\, \omega_y\, \varphi(z) = \sum_{\ell=0}^{\infty} e_\ell(y) \sum_{i=-\infty}^{\infty} e_i(x) z^{i-\ell}$$

$$= \prod_{n=1}^{\infty} \left(1 + y_n\, z^{-1}\right) \prod_{n=1}^{\infty} \left(1 + x_n\, z\right). \tag{2.5}$$

We remark that $\omega_y R_n(x,y)$ also equals a Toeplitz determinant. The $n \to \infty$ limit of this identity is, by an application of the strong Szegö theorem, the *dual* Cauchy identity

$$\sum_{\lambda\in\mathcal{P}} s_\lambda(x) s_{\lambda'}(y) = \prod_{i,j} (1 + x_i y_j).$$

5. Specializations

If R is a commutative $\mathbf{Q}$-algebra with identity, then a *specialization* of the ring Λ is a homomorphism $\psi : \Lambda \to R$. We always assume that $\psi(1) = 1$.

(i) Exponential specialization. If $p_n = \sum_i x_i^n$, the power sum symmetric functions, then the exponential specialization is determined by

$$ex(p_n) = t\delta_{1n}.$$

(Recall that the p_n's are algebraically independent generators of Λ. This homomorphism ex is denoted by θ in Gessel.)

The fundamental property of this homomorphism is for any symmetric function f

$$ex(f) = \sum_{n\geq 0} [x_1x_2\cdots x_n]f \, \frac{t^n}{n!}$$

where $[x_1x_2\cdots x_n]f$ denotes the coefficient of $x_1x_2\cdots x_n$ in f. Thus if $\lambda \vdash N$ and s_λ is the Schur function, then

$$ex(s_\lambda) = f^\lambda \, \frac{t^N}{N!}$$

where f^λ is the number of standard Young tableaux of shape λ. Hence

$$ex_x \, ex_y R_n(x,y) = \sum_{N=0}^{\infty} \sum_{\substack{\lambda\vdash N \\ \ell(\lambda)\leq n}} \left(f^\lambda\right)^2 \frac{t^{2N}}{(N!)^2}.$$

By the Robinson-Schensted-Knuth (RSK) bijection [31, 32, 23]

$$\sum_{\substack{\lambda\vdash N \\ \ell(\lambda)\leq n}} \left(f^\lambda\right)^2$$

equals the number of permutations σ on N letters such that $\ell_N(\sigma)$, the length of the longest increasing subsequence in σ, is at most n. Hence if each such permutation has probability $1/N!$ we have

$$F_P(n;N) := \mathrm{Prob}\left(\ell_N(\sigma) \leq n\right) = \frac{1}{N!} \sum_{\substack{\lambda\vdash N \\ \ell(\lambda)\leq n}} \left(f^\lambda\right)^2.$$

Thus we know that its generating function is given by

$$\sum_{N=0}^{\infty} F_P(n;N) \, \frac{t^{2N}}{N!} = ex_x \, ex_y R_n(x,y).$$

Gessel's theorem tells us that $R_n(x,y)$ is the $n\times n$ Toeplitz determinant with symbol $\varphi(z)$ given by (2.4). This may be written

$$\varphi(z) = \sum_{r\geq 0} h_r(y) z^{-r} \sum_{s\geq 0} h_s(x) z^s.$$

The important observation is that since ex is a homomorphism, $ex_x\, ex_y R_n(x,y)$ is the Toeplitz determinant with symbol

$$f_P(z) := ex_x\, ex_y(\varphi(z)) = \sum_{r\geq 0} \frac{t^r}{r!} z^{-r} \sum_{s\geq 0} \frac{t^s}{s!} z^s = e^{t\,(z+z^{-1})}.$$

This is precisely (1.1) after changing t to $\sqrt{t}$.

(ii) Principal specializations. The *principal specialization* of order n of g is defined by

$$ps_n(g) = g(1, q, q^2, \ldots, q^{n-1}, 0, 0, \ldots).$$

(Thus we replace x_i by q^i if $i < n$ and by 0 itherwise. If we let $n \to \infty$ we obtain the *stable principal specialization* of f.) Setting $q = 1$ in ps_n gives

$$ps_n^1(g) = g(1, 1, \ldots, 1, 0, 0, \ldots)$$

where 1 appears n times.

Observe that

$$ps_k^1(s_\lambda) = d_\lambda(k)$$

where $d_\lambda(k)$ is the number of semistandard Young tableaux of shape λ that can be formed from an alphabet of k letters. (Recall that a semistandard tableau is weakly increasing across rows and strictly increasing down columns; a standard tableau is strictly increasing across rows.) This is most easily seen from the combinatorial definition of the Schur functions.

Applying the homomorphisms ps_n^1 and ex to R_n gives

$$(ps_k^1)_x\, ex_y\, R_n(x,y) = \sum_{N=0}^{\infty} \left(\sum_{\substack{\lambda\vdash N \\ \ell(\lambda)\leq n}} d_\lambda(k)\, f^\lambda \right) \frac{t^N}{N!}. \tag{2.6}$$

The RSK correspondence associates to each word w of length N formed from an alphabet of k letters a pair of tableaux, (P, Q). Here the P are semistandard Young tableaux of shape $\lambda \vdash N$ made from the alphabet $\{1, 2, \ldots, k\}$ and the Q are standard Young tableaux of shape $\lambda \vdash N$ on the numbers $\{1, 2, \ldots, N\}$. Thus

$$\sum_{\substack{\lambda\vdash N \\ \ell(\lambda)\leq n}} d_\lambda(k)\, f^\lambda$$

counts the number of words w of length N with strictly decreasing subsequences all of length less than or equal to n. Obviously,

$$\sum_{\lambda\vdash N} d_\lambda(k)\, f^\lambda = k^N.$$

The symbol of the Toeplitz determinant that equals the generating function (2.6) is

$$f_D(z) = f_D(z;\, t) := (ps_k^1)_x ex_y(\varphi(z)) = \prod_{i=1}^{k} (1-z)^{-1} \sum_{r\geq 0} \frac{t^r}{r!} z^{-r} = \frac{e^{t/z}}{(1-z)^k}.$$

Hence we have shown that if $\ell_N^D(w)$ denotes the length of the longest *strictly decreasing* subsequence in word w, and if each word of length N is assigned probability $1/k^N$, then the generating function of the distribution function

$$F_D(n;k,N) := \mathrm{Prob}_k\left(\ell_N^D(w) \le n\right)$$

is given by the Toeplitz determinant with symbol f_D:

$$\sum_{N=0}^{\infty} F_D(n;k,N)\,\frac{(kt)^N}{N!} = \det\left(T_n(f_D(z;\,t))\right).$$

Recalling (1.4) we see that this is

$$G_D(n;k,kt) = \det\left(T_n(f_D(z;\,t))\right). \tag{2.7}$$

Since, under general conditions, changing the symbol of a Toeplitz matrix from $f(z)$ to $f(az)$ is a similarity transformation, the associated Toeplitz determinant does not change. Therefore the symbol of the Toeplitz determinant in (2.7) may be replaced by

$$f(\sqrt{t}z/k;\,t/k) = \frac{e^{\sqrt{t}/z}}{(1-\sqrt{t}z/k)^k}$$

whose $k\to\infty$ limit is $e^{\sqrt{t}\,(z+z^{-1})}$. This shows that for fixed N,

$$\lim_{k\to\infty} F_D(n;k,N) = F_P(n;N).$$

Again, this is intuitively clear since as the size of the alphabet approaches infinity, any random word of length N is very likely a permutation. (This also uses the fact that the distribution of the length of the longest decreasing subsequence of a random permutation is the same as the distribution of the length of the longest increasing subsequence.)

Finally we apply the same specialization $(ps_k^1)_x\, ex_y$ to the dual version of Gessel's theorem. We see that (2.6) is replaced by

$$(ps_k^1)_x\, ex_y\, \tilde{R}_n(x,y) = \sum_{N=0}^{\infty}\left(\sum_{\substack{\lambda_1\le n\\ \lambda\in\mathcal{P}}} d_\lambda(k) f^\lambda\right)\frac{t^N}{N!}. \tag{2.8}$$

(We used here the fact that $f^{\lambda'} = f^\lambda$, which follows immediately from the hook length formula for f^λ.) Thus we obtain the generating function for the distribution of the length of the longest *weakly increasing* subsequence. Using (2.5) we find that the symbol of the Toeplitz determinant that gives (2.8) is

$$f_I(z) = f_I(z;\,t) := (ps_k^1)_x\, ex_y \sum_{r\ge 0} e_r(x) z^r \sum_{s\ge 0} e_s(y) z^{-s}$$

$$= (ps_k^1)_x\, ex_x \prod_j (1+x_j z) \sum_{s\ge 0} e_s(y) z^{-s} = (1+z)^k\, e^{t/z}.$$

To summarize, we have shown that if $\ell_N^I(w)$ denotes the length of the longest weakly increasing subsequence in word w of length N, and if each such word has probability $1/k^N$, then the generating function of the distribution function

$$F_I(n;k,N) := \mathrm{Prob}_k\left(\ell_N^I(w) \le n\right)$$

is given by the Toeplitz determinant with symbol f_I. Precisely,

$$G_I(n;k,kt) = \sum_{N=0}^{\infty} F_I(n;k,N)\,\frac{(kt)^N}{N!} = \det\left(T_n(f_I(z;\,t))\right). \tag{2.9}$$

The same $k \to \infty$ remarks hold here as in the strictly decreasing case.

Relations (2.7) and (2.9) are the assertions of Theorem 1.

III. Recursion and differentiation formulas

1. Universal recursion relations

In this section f will be an arbitrary function, with Fourier coefficients f_i and associated $n \times n$ Toeplitz matrix

$$T_n(f) = (f_{i-j}), \qquad (i,\,j = 0,\cdots,n-1).$$

We assume $T_n(f)$ is invertible and obtain several relations connecting various inner products involving $T_n(f)^{-1}$. Most of these relations actually appeared in [37]. There our $T_n(f)$ was symmetric and unfortunately some of the relations derived in [37] used this fact. Since this does not happen here, everything has to be modified for the more general case. A reader of the earlier article will find familiar much of what we now do.

We introduce the n-vectors

$$\delta^+ = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \\ 0 \end{pmatrix}, \quad \delta^- = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}, \quad f^+ = \begin{pmatrix} f_1 \\ f_2 \\ \vdots \\ f_{n-1} \\ f_n \end{pmatrix}, \quad f^- = \begin{pmatrix} f_n \\ f_{n-1} \\ \vdots \\ f_2 \\ f_1 \end{pmatrix}$$

and define $\widetilde{f}$ by $\widetilde{f}(z) = f(z^{-1})$, so that $T_n(\widetilde{f})$ is the transpose of $T_n(f)$. We write

$$\Lambda = T_n(z^{-1}), \quad \Lambda' = T_n(z).$$

Thus Λ is the backward shift and Λ' is the forward shift. It is easy to see that

$$T_n(z^{-1}\,f) = T_n(f)\,\Lambda + f^+ \otimes \delta^+ = \Lambda\,T_n(f) + \delta^- \otimes f^-, \tag{3.1}$$

$$T_n(z\,f) = T_n(f)\,\Lambda' + \widetilde{f}^- \otimes \delta^- = \Lambda'\,T_n(f) + \delta^+ \otimes \widetilde{f}^+. \tag{3.2}$$

These identities explain why the vectors $f^\pm$ and $\widetilde{f}^\pm$ arise.

The inner products involving $T_n(f)^{-1}$ are

$$U_n^{\pm} = (T_n(f)^{-1} f^+,\, \delta^{\pm}), \quad V_n^{\pm} = (T_n(f)^{-1} \delta^+,\, \delta^{\pm}).$$

If one of these quantities defined in terms of f is given a tilde, then f is to be replaced by $\widetilde{f}$ everywhere in its definition. Thus, for example,

$$\widetilde{U}_n^+ = (T_n(\widetilde{f})^{-1} \widetilde{f}^+,\, \delta^+).$$

Note that $\widetilde{V}_n^+ = V_n^+ = D_n/D_{n+1}$, where

$$D_n = D_n(f) = \det\, T_n(f).$$

Some other inner products may be expressed in terms of these using the isometry that reverses the order of the components of a vector (and replaces a Toeplitz matrix by its transpose). Thus, for example,

$$(T_n(\widetilde{f})^{-1} f^-,\, \delta^+) = (T_n(f)^{-1} f^+,\, \delta^-) = U_n^-. \tag{3.3}$$

We shall use this isometry from time to time below without comment.

The basis for all the universal relations we shall obtain is the following formula for the inverse of a 2×2 block matrix:

$$\begin{pmatrix} A & B \\ C & D \end{pmatrix}^{-1} = \begin{pmatrix} (A - BD^{-1}C)^{-1} & \times \\ \times & \times \end{pmatrix}. \tag{3.4}$$

Here we assume A and D are square and the various inverses exist. Only one block of the inverse is displayed and the formula shows that $A - BD^{-1}C$ equals the inverse of this block of the inverse matrix.

We apply (3.4) first to the $(n+1) \times (n+1)$ matrix

$$\begin{pmatrix} 0 & 0 & \cdots & 1 \\ f_1 & f_0 & \cdots & f_{-n+1} \\ \vdots & \vdots & \cdots & \vdots \\ f_n & f_{n-1} & \cdots & f_0 \end{pmatrix},$$

with $A = (0)$, $D = T_n(f)$, $B = (0\ \cdots\ 0\ 1)$, $C = f^+$. In this case $A - BD^{-1}C = -(T_n(f)^{-1} f^+,\, \delta^-) = -U_n^-$. This equals the reciprocal of the upper-left entry of the inverse matrix, which in turn equals $(-1)^n$ times the lower-left $n \times n$ subdeterminant divided by D_n. Replacing the first row by $(f_0\ f_{-1}\ \cdots\ f_{-n})$ gives the matrix

$$\begin{pmatrix} f_0 & f_{-1} & \cdots & f_{-n} \\ f_1 & f_0 & \cdots & f_{-n+1} \\ \vdots & \vdots & \cdots & \vdots \\ f_n & f_{n-1} & \cdots & f_0 \end{pmatrix} = T_{n+1}(f). \tag{3.5}$$

The lower-left entry of its inverse equals on the one hand $(T_n(f)^{-1}\delta^+,\,\delta^-)=V_{n+1}^-$ and on the other hand $(-1)^n$ times the same subdeterminant as arose above divided by U_{n+1}^-. This gives the identity

$$-\,U_n^-=V_{n+1}^-\,\frac{D_{n+1}}{D_n}=\frac{V_{n+1}^-}{V_{n+1}^+}. \tag{3.6}$$

If we take A to be the upper-left corner of (3.5) and D the complementary $T_n(f)$ then $C=f^+$ and $B=(f_{-1}\ \cdots\ f_{-n})$, and we deduce that

$$f_0-(T_n(f)^{-1}f^+,\,\widetilde{f}^+)=\frac{1}{V_{n+1}^+}. \tag{3.7}$$

Next we consider

$$\begin{pmatrix} f_0 & f_{-1} & \cdots & f_{-n} & f_{-n-1}\\ f_1 & f_0 & \cdots & f_{-n+1} & f_{-n}\\ \vdots & \vdots & \cdots & \vdots & \vdots\\ f_n & f_{n-1} & \cdots & f_0 & f_{-1}\\ f_{n+1} & f_n & \cdots & f_1 & f_0 \end{pmatrix}=T_{n+2}(f).$$

We apply to this an obvious modification of (3.4), where A is the 2×2 matrix consisting of the four corners of the large matrix, D is the central $T_n(f)$, C consists of the two columns f^+ and $\widetilde{f}^-$ and B consists of the rows which are the transposes of $\widetilde{f}^+$ and f^-. Then we find

$$A-BD^{-1}C=\begin{pmatrix} f_0-(T_n(f)^{-1}f^+,\,\widetilde{f}^+) & f_{-n-1}-(T_n(f)^{-1}\widetilde{f}^-,\,\widetilde{f}^+)\\ \\ f_{n+1}-(T_n(f)^{-1}f^+,\,f^-) & f_0-(T_n(f)^{-1}\widetilde{f}^-,\,f^-) \end{pmatrix}$$

and our formula tells us that this is the inverse of

$$\begin{pmatrix} V_{n+2}^+ & \widetilde{V}_{n+2}^-\\ \\ V_{n+2}^- & V_{n+2}^+ \end{pmatrix}.$$

This gives the two formulas

$$f_0-(T_n(f)^{-1}f^+,\,\widetilde{f}^+)=\frac{V_{n+2}^+}{{V_{n+2}^+}^2-V_{n+2}^-\widetilde{V}_{n+2}^-},\quad f_{n+1}-(T_n(f)^{-1}f^+,\,f^-)=\frac{-V_{n+2}^-}{{V_{n+2}^+}^2-V_{n+2}^-\widetilde{V}_{n+2}^-}.$$

Comparing the first with (3.7) we see that

$${V_{n+2}^+}^2-V_{n+2}^-\widetilde{V}_{n+2}^-=V_{n+1}^+\,V_{n+2}^+, \tag{3.8}$$

and therefore that the preceding relations can be written

$$f_0-(T_n(f)^{-1}f^+,\,\widetilde{f}^+)=\frac{1}{V_{n+1}^+},\quad f_{n+1}-(T_n(f)^{-1}f^+,\,f^-)=-\frac{1}{V_{n+1}^+}\,\frac{V_{n+2}^-}{V_{n+2}^+}. \tag{3.9}$$

Notice that (3.6) and (3.8) give

$$1 - U_n^- \, \widetilde{U}_n^- = \frac{V_n^+}{V_{n+1}^+}. \tag{3.10}$$

Next, we apply (3.4) to the matrix

$$\begin{pmatrix} 0 & 1 & \cdots & 0 \\ f_1 & f_0 & \cdots & f_{-n+1} \\ \vdots & \vdots & \cdots & \vdots \\ f_n & f_n & \cdots & f_0 \end{pmatrix}.$$

Now we have $A = (0)$, $\;D = T_n(f)$, $B = (1 \;\cdots\; 0 \; 0)$, $\;C = f^+$ and so $A - BD^{-1}C = -(T_n(f)^{-1} \, f^+, \, \delta^+) = -U_n^+$. Therefore this equals the inverse of the $0,0$ entry of the inverse of the matrix, which in turn equals its determinant divided by D_n. But its determinant equals $-K$, where K is the $0,1$ cofactor. Thus

$$U_n^+ = \frac{K}{D_n}.$$

Now look at the matrix (3.5) and consider the $1,0$ entry of its inverse. It equals on the one hand $(T_{n+1}(f))^{-1} \, \delta^+, \, \Lambda'\delta^+)$ and on the other hand $-K/D_{n+1}$. This gives the identity

$$U_n^+ = -\frac{D_{n+1}}{D_n} \, (T_{n+1}(f)^{-1} \, \delta^+, \, \Lambda'\delta^+) = -\frac{1}{V_{n+1}^+} \, (T_{n+1}(f)^{-1} \, \delta^+, \, \Lambda'\delta^+). \tag{3.11}$$

To evaluate the inner product on the right side we for the moment replace $n+1$ by n so that we can apply earlier formulas. The inner product becomes

$$(T_n(f)^{-1}\delta^+, \, \Lambda'\delta^+) = (\Lambda \, T_n(f)^{-1}\delta^+, \, \delta^+).$$

Multiplying the second identity of (3.1) left and right by $T_n(f)^{-1}$ gives

$$\Lambda \, T_n(f)^{-1} + T_n(f)^{-1} f^+ \otimes T_n(\widetilde{f})^{-1}\delta^+ = T_n(f)^{-1} \, \Lambda + T_n(f)^{-1}\delta^- \otimes T_n(\widetilde{f})^{-1} f^-. \tag{3.12}$$

Applying this to δ^+ (observing that $\Lambda\delta^+ = 0$) and taking the inner product with δ^+ we find that

$$(\Lambda \, T_n(f)^{-1}\delta^+, \, \delta^+) = \widetilde{V}_n^- \, U_n^- - U_n^+ \, V_n^+.$$

Here we used (3.3). Replacing n by $n+1$, we see that (3.11) becomes

$$U_n^+ = -\frac{\widetilde{V}_{n+1}^-}{V_{n+1}^+} \, U_{n+1}^- + U_{n+1}^+.$$

This gives

$$U_n^+ - U_{n+1}^+ = \widetilde{U}_n^- \, U_{n+1}^- \tag{3.13}$$

by (3.6).

Next we apply (3.4) to the $n \times n$ matrix

$$\begin{pmatrix} 0 & f_{-1} & \cdots & f_{-n} \\ 0 & f_0 & \cdots & f_{-n+1} \\ \vdots & \vdots & \cdots & \vdots \\ 1 & f_{n-2} & \cdots & f_{-1} \\ 0 & f_{n-1} & \cdots & f_0 \end{pmatrix}.$$

The formula says that the inverse of the upper-left entry of its inverse equals

$$-(T_n(f)^{-1}\,\Lambda\delta^-,\ \tilde{f}^+).$$

This inverse also equals $(-1)^{n-1}\,D_n/K$ where K is the $n-1$, 0 cofactor of the matrix (3.5). But $(T_{n+1}(f)^{-1}\Lambda\delta^-,\ \delta^+)$, the 0, $n-1$ entry of the inverse of (5), equals $(-1)^{n-1}\,K/D_{n+1}$. We have shown

$$\frac{D_{n+1}}{D_n}\,(T_{n+1}(f)^{-1}\,\Lambda\,\delta^-,\ \delta^+) = -(T_n(f)^{-1}\,\Lambda\,\delta^-,\ \tilde{f}^+).$$

This may be written

$$(\delta^+,\ \Lambda\,T_{n+1}(f)^{-1}\,\delta^-) = -V_{n+1}\,(\delta^+,\ \Lambda\,T_n(f)^{-1}\,\tilde{f}^-).$$

To compute the left side we use the fact that

$$\Lambda\,\tilde{f}^- = \begin{pmatrix} f_{-n+1} \\ \vdots \\ f_{-1} \\ 0 \end{pmatrix} = T_n(f)\,\delta^- - f_0\,\delta^-.$$

Hence

$$T_n(f)^{-1}\,\Lambda\,\tilde{f}^- = \delta^- - f_0\;T_n(f)^{-1}\delta^-.$$

Therefore applying (3.12) to $\tilde{f}^-$ gives

$$\Lambda\,T_n(f)^{-1}\,\tilde{f}^- = \delta^- - f_0\;T_n(f)^{-1}\delta^-$$
$$+(T_n(\tilde{f})^{-1}f^-,\ \tilde{f}^-)\;T_n(f)^{-1}\delta^- - (T_n(\tilde{f})^{-1}\,\delta^+,\ \tilde{f}^-)\;T_n(f)^{-1}f^+.$$

Taking inner products with δ^+ we find

$$(\delta^+,\ \Lambda\,T_n(f)^{-1}\,\tilde{f}^-) = -\Big(f_0 - (T_n(\tilde{f})^{-1}f^-,\ \tilde{f}^-)\Big)\;\tilde{V}_n^- - \tilde{U}_n^-\,U_n^+.$$

But (3.9) gives

$$f_0 - (T_n(\tilde{f})^{-1}f^-,\ \tilde{f}^-) = f_0 - (T_n(f)^{-1}f^+,\ \tilde{f}^+) = \frac{1}{V_{n+1}^+}.$$

Hence

$$(\delta^+,\ \Lambda\,T_n(f)^{-1}\,\tilde{f}^-) = -\frac{\tilde{V}_n^-}{V_{n+1}^+} - \tilde{U}_n^-\,U_n^+,$$

$$(\delta^+,\ \Lambda\,T_{n+1}(f)^{-1}\,\delta^-) = \tilde{V}_n^- + \tilde{U}_n^-\,U_n^+\,V_{n+1}^+. \tag{3.14}$$

2. Nonuniversal recursion relations

Here we restrict to our symbol

$$f_I(z) = e^{t/z}\,(1+z)^k$$

but write f instead of f_I for notational convenience. We shall obtain relations which follow from the representation

$$f_j = \frac{1}{2\pi i}\int e^{t/z}\,(1+z)^k\,z^{-j-1}\,dz$$

upon integrating by parts. The fact

$$\frac{1}{2\pi i}\int \frac{d}{dz}\left\{e^{t/z}\,(1+z)^{k+1}\,z^{-j}\right\}\,dz = 0$$

gives

$$-t\,(f_{j+1}+f_j) + (k+1)\,f_{j-1} - j\,(f_j + f_{j-1}) = 0,$$

$$(j+t)\,f_j + (j-k-1)\,f_{j-1} + t\,f_{j+1} = 0.$$

Replacing j by $i-j$ we obtain the i,j entry of a matrix identity:

$$(M+t)\,T_n(f) - T_n(f)\,M + (M-k-1)\,T_n(zf) - T_n(zf)\,M + t\,T(z^{-1}f) = 0.$$

Here M denotes the diagonal matrix with diagonal entries $1,\,2,\cdots,n$. We use the identities (3.1) and (3.2) to write this as

$$(M+t)\,T_n(f) - T_n(f)\,M + (M-k-1)\,(\Lambda'\,T_n(f) + \delta^+\otimes\widetilde{f}^+)$$

$$-(T_n(f)\,\Lambda' + \widetilde{f}^-\otimes\delta^-)M + t\,(T_n(f)\,\Lambda + f^+\otimes\delta^+) = 0,$$

$$(M+t)\,T_n(f) - T_n(f)\,M + (1-k)\,\Lambda'\,T_n(f) - T_n(f)\,\Lambda' + t\,T_n(f)\,\Lambda$$

$$-k\,\delta^+\otimes\widetilde{f}^+ - n\,\widetilde{f}^-\otimes\delta^- + t\,f^+\otimes\delta^+ = 0.$$

We multiply this left and right by $T_n(f)^{-1}$, obtaining

$$T_n(f)^{-1}\,(M+t) - M\,T_n(f)^{-1} + T_n(f)^{-1}\,(M-k-1)\,\Lambda' - \Lambda'\,M\,T_n(f)^{-1} + t\,\Lambda\,T_n(f)^{-1}$$

$$-k\,T_n(f)^{-1}\delta^+\otimes T_n(\widetilde{f})^{-1}\widetilde{f}^+ - n\,T_n(f)^{-1}\widetilde{f}^-\otimes T_n(\widetilde{f})^{-1}\delta^- + t\,T_n(f)^{-1}f^+\otimes T_n(\widetilde{f})^{-1}\delta^+ = 0. \quad (3.15)$$

This is the basic matrix identity. Applying this matrix identity to $\delta^\pm$ and taking inner products with $\delta^\pm$ gives identities for scalar quantities. We shall need three of the four.

First we apply (3.15) to δ^+ and take the inner product with δ^+. If we recall our definitions and the fact $\Lambda\delta^+ = 0$ we obtain

$$t\,V_n^+ + (1-k)\,(T_n(f)^{-1}\,\Lambda'\,\delta^+,\,\delta^+) + t\,(\Lambda\,T_n(f)^{-1}\delta^+,\,\delta^+)$$

$$-\,k\,V_n^+\,\widetilde{U}_n^+ - n\,(T_n(f)^{-1}\widetilde{f}^-,\,\delta^+)\,V_n^- + t\,U_n^+V_n^+ = 0. \quad (3.16)$$

But

$$(T_n(f)^{-1}\widetilde{f}^-,\,\delta^+) = (T_n(\widetilde{f})^{-1}\widetilde{f}^+,\,\delta^-) = \widetilde{U}_n^-,$$

so (3.6) gives

$$V_n^- = -V_n^+ \, U_{n-1}^-.$$

Thus

$$(T_n(f)^{-1} \widetilde{f}^-, \, \delta^+) \, V_n^- = -V_n^+ \, \widetilde{U}_n^- \, U_{n-1}^-.$$

Next, by (3.11) we have

$$(T_n(f)^{-1} \, \Lambda' \, \delta^+, \, \delta^+) = -V_n^+ \, \widetilde{U}_{n-1}^+, \qquad (\Lambda \, T_n(f)^{-1} \delta^+, \, \delta^+) = -V_n^+ \, U_{n-1}^+.$$

Substituting these relations into (3.16) and dividing by V_n^+ gives

$$t + (k-1) \, \widetilde{U}_{n-1}^+ - t \, U_{n-1}^+ - k \, \widetilde{U}_n^+ + n \, \widetilde{U}_n^- \, U_{n-1}^- + t \, U_n^+ = 0.$$

But (3.13) gives $\widetilde{U}_n^- \, U_{n-1}^- = \widetilde{U}_{n-1}^+ - \widetilde{U}_n^+$, from which we see that the above becomes

$$t + (k+n-1) \, \widetilde{U}_{n-1}^+ - (k+n) \, \widetilde{U}_n^+ + t \, (U_n^+ - U_{n-1}^+) = 0.$$

The derivation holds also for $n = 1$ when one defines $U_0^+ = \widetilde{U}_0^+ = 0$, as one can easily check. Therefore summing over n gives

$$n \, t - (k+n) \, \widetilde{U}_n^+ + t \, U_n^+ = 0. \tag{3.17}$$

For the next relation we apply (3.15) to δ^- and take the inner product with δ^+ to obtain

$$(n+t-1) \, \widetilde{V}_n^- + t \, (\Lambda \, T_n(f)^{-1} \, \delta^-, \, \delta^+) - k \, V_n^+ \, \widetilde{U}_n^- - n \, \widetilde{U}_n^- \, V_n^+ + t \, U_n^+ \, \widetilde{V}_n^- = 0.$$

If we applying (3.14), divide by V_n^+ and use (3.6) we obtain

$$-(n+t-1) \, \widetilde{U}_{n-1}^- + t \left(\frac{\widetilde{V}_{n-1}^-}{V_n^+} + U_{n-1}^+ \, \widetilde{U}_{n-1}^- \right) - (k+n) \, \widetilde{U}_n^- - t \, U_n^+ \, \widetilde{U}_{n-1}^- = 0.$$

But

$$\frac{\widetilde{V}_{n-1}^-}{V_n^+} = \frac{\widetilde{V}_{n-1}^-}{V_{n-1}^+} \, \frac{\widetilde{V}_{n-1}^+}{V_n^+} = -\widetilde{U}_{n-2}^- \, (1 - U_{n-1}^- \, \widetilde{U}_{n-1}^-),$$

by (3.6) and (3.10). Replacing n by $n+1$ and introducing the function

$$\Phi_n = 1 - U_n^- \, \widetilde{U}_n^-$$

we obtain

$$-(t+n) \, \widetilde{U}_n^- + t \, (U_n^+ \, \widetilde{U}_n^- - \widetilde{U}_{n-1}^- \, \Phi_n) - (k+n+1) \, \widetilde{U}_{n+1}^- - t \, U_{n+1}^+ \, \widetilde{U}_n^- = 0.$$

By (3.13) this may be written

$$-\,(t+n) \, \widetilde{U}_n^- + t \, (\widetilde{U}_n^-)^2 \, U_{n+1}^- - t \, \widetilde{U}_{n-1}^- \, \Phi_n - (k+n+1) \, \widetilde{U}_{n+1}^- = 0. \tag{3.18}$$

Finally we we apply (3.15) to δ^- and take the inner product with δ^- to obtain

$$n \, V_n^+ - (n-1) \, (T_n(f)^{-1} \delta^-, \, \Lambda \, \delta^-) - k \, V_n^- \, \widetilde{U}_n^- - n \, \widetilde{U}_n^+ \, V_n^+ + t \, U_n^- \, \widetilde{V}_n^- = 0.$$

From (3.11) we see that $(T_n(f)^{-1}\delta^-,\,\Lambda\,\delta^-) = -\widetilde{U}_{n-1}^+\,V_n^+$. Substituting this into the above identity, dividing by V_n and using (3.6) gives

$$t + (n-1)\,\widetilde{U}_{n-1}^+ + k\,U_{n-1}^-\,\widetilde{U}_n^- - n\,\widetilde{U}_n^+ - t\,U_n^-\,\widetilde{U}_{n-1}^- = 0. \tag{3.19}$$

Using (3.13) we can write

$$k\,U_{n-1}^-\,\widetilde{U}_n^- = n\,U_{n-1}^-\,\widetilde{U}_n^- + (k-n)\,(\widetilde{U}_{n-1}^+ - \widetilde{U}_n^+).$$

Substituting this into the preceding and replacing n by $n+1$ give

$$t + (k-1)\,\widetilde{U}_n^+ + (n+1)\,U_n^-\,\widetilde{U}_{n+1}^- - k\,\widetilde{U}_{n+1}^+ - t\,U_{n+1}^-\,\widetilde{U}_n^- = 0. \tag{3.20}$$

Another identity can be gotten by using (3.13) to replace $n\,(\widetilde{U}_{n-1}^- - \widetilde{U}_n^-)$, which we see in (3.19), by $n\,U_{n-1}^-\,\widetilde{U}_n^-$. This gives

$$t - \widetilde{U}_{n-1}^+ + (k+n)\,U_{n-1}^-\,\widetilde{U}_n^- - t\,U_n^-\,\widetilde{U}_{n-1}^- = 0.$$

Replacing n by $n+1$, multiplying by $\widetilde{U}_n^-$ and adding to (3.18) gives

$$n\,\widetilde{U}_n^- + \widetilde{U}_n^-\,\widetilde{U}_n^+ + \Phi_n\,\Big((k+n+1)\,\widetilde{U}_{n+1}^- + t\,\widetilde{U}_{n-1}^-\Big) = 0. \tag{3.21}$$

3. Differentiation formulas

We continue to take $f(z) = e^{t/z}\,(1+z)^k$ and write $D_n(t)$ for $D_n(f)$ Since $df/dt = z^{-1}f$ we have

$$\frac{d}{dt}\log\,D_n(t) = \operatorname{tr} T_n(f)^{-1}\,T_n(z^{-1}f) = \operatorname{tr}\,\Big(\Lambda + T_n(f)^{-1}f^+\otimes\delta^+\Big)$$

by (3.1), and so

$$\frac{d}{dt}\log\,D_n(t) = U_n^+, \tag{3.22}$$

which is why this quantity arises. Others will arise from further differentiation.

We use

$$\frac{d}{dt}T_n(f)^{-1} = -T_n(f)^{-1}\,T_n(z^{-1}f)\,T_n(f)^{-1} = -\Lambda\,T_n(f)^{-1} - T_n(f)^{-1}f^+\otimes T_n(\widetilde{f})^{-1}\delta^+,$$

$$\frac{df^+}{dt} = (z^{-1}f)^+ = \Lambda\,f^+ + f_{n+1}\,\delta^-.$$

Hence

$$\begin{aligned}\frac{dU_n^+}{dt} &= -(\Lambda\,T_n(f)^{-1}f^+,\,\delta^+) - (T_n(f)^{-1}f^+,\,\delta^+)\,(T_n(\widetilde{f})^{-1}\delta^+,\,f^+)\\ &\quad +(T_n(f)^{-1}\,\Lambda\,f^+,\,\delta^+) + f_{n+1}\,(T_n(f)^{-1}\delta^-,\,\delta^+).\end{aligned}$$

By (3.12) the two terms involving Λ combine to give

$$(T_n(f)^{-1}f^+,\,\delta^+)\,(T_n(\widetilde{f})^{-1}\delta^+,\,f^+) - (T_n(f)^{-1}\delta^-,\,\delta^+)\,(T_n(\widetilde{f})^{-1}f^-,\,f^+).$$

Using this we find that the preceding simplifies to

$$-\widetilde{V}_n^-\,(T_n(\widetilde{f})^{-1}f^-,\,f^+)+f_{n+1}\,\widetilde{V}_n^-=-\widetilde{V}_n^-\,\frac{1}{V_{n+1}^+}\,\frac{V_{n+2}^-}{V_{n+2}^+},$$

by the second part of (3.9). This equals

$$-\frac{\widetilde{V}_n^-}{V_n^+}\,\frac{V_n^+}{V_{n+1}^+}\,\frac{\widetilde{V}_{n+2}^-}{\widetilde{V}_{n+2}^+}=-\widetilde{U}_{n-1}^-\,(1-U_n^-\,\widetilde{U}_n^-)\,U_{n+1}^-,$$

by (3.6) and (3.10). We have shown

$$\frac{dU_n^+}{dt}=-\Phi_n\,\widetilde{U}_{n-1}^-\,U_{n+1}^-. \tag{3.23}$$

In completely analogous fashion (we spare the reader the details) we compute

$$\frac{dU_n^-}{dt}=-V_n^+\,(T_n(\widetilde{f})^{-1}f^-,\,f^+)+f_{n+1}\,V_n^+,$$

and using again the second part of (3.9), (3.6) and (3.10) we find that

$$\frac{dU_n^-}{dt}=\Phi_n\,U_{n+1}^-. \tag{3.24}$$

To find formulas for the derivatives of $\widetilde{U}_n^{\pm}$ we use

$$\frac{d}{dt}T_n(\widetilde{f})^{-1}=-T_n(f)^{-1}\,T_n(zf)\,T_n(f)^{-1}=-\Lambda'\,T_n(f)^{-1}+T_n(f)^{-1}\widetilde{f}^-\otimes T_n(\widetilde{f})^{-1}\delta^-,$$

$$\frac{d\widetilde{f}^+}{dt}=(z\widetilde{f})^+=\Lambda'\,f^++f_0\,\delta^+.$$

At the appropriate points in the computations we use the analogue of (3.12) with Λ' instead of Λ, and the first part of (3.9) rather than the second. Again we spare the reader the details. The results are somewhat simpler:

$$\frac{d\widetilde{U}_n^+}{dt}=\Phi_n, \tag{3.25}$$

$$\frac{d\widetilde{U}_n^-}{dt}=-\Phi_n\,\widetilde{U}_{n-1}^-. \tag{3.26}$$

Observe that from (3.22), (3.17), and (3.25) we have

$$\frac{d}{dt}t\,\frac{d}{dt}\log\,D_n(t)=(k+n)\,\Phi_n-n.$$

This gives the representation

$$\log\,D_n(t)=(k+n)\,\int_0^t\log(t/t')\,\Phi_n(t')\,dt'-nt.$$

IV. Painlevé V and the Laguerre ensemble

1. Derivation of the differential equation

We begin by showing how differentiation formulas (3.24) and (3.26) have analogues in which only indices n and $n-1$ appear on the right side of (3.24) and only indices n and $n+1$ appear on the right side of (3.26).

Solve (3.17) for U^-_{n+1}. The solution involves $\widetilde{U}^-_{n+1}$, which we can solve for in (3.20). Thus U^-_{n+1}, and so dU^-_n/dt, can be expressed in terms of quantities with indices n or $n-1$. To obtain a differentiation formula for $\widetilde{U}^-_n$ that involves only n and $n+1$ simply solve (3.20) for $\widetilde{U}^-_{n-1}$. The results of this are

$$\frac{dU^-_n}{dt} = -\frac{n}{t}\,U^-_n + \frac{1}{t}\,\left(\widetilde{U}^+_n - t\Phi_n\right)\frac{U^-_n}{\Phi_n - 1} + \frac{\Phi_n}{\Phi_n - 1}\widetilde{U}^-_{n-1}{U^-_n}^2, \tag{4.1}$$

$$\frac{d\widetilde{U}^-_n}{dt} = \frac{n}{t}\,\widetilde{U}^-_n + \frac{1}{t}\,\widetilde{U}^+_n\widetilde{U}^-_n + \frac{1}{t}(k+1+n)\,\Phi_n\,\widetilde{U}^-_{n+1}. \tag{4.2}$$

Now compute Φ'_n using (3.24) and (4.2). The result can be put in the form

$$\frac{\Phi'_n}{\Phi_n} - \frac{n}{t}\frac{\Phi_n - 1}{\Phi_n} - \frac{1}{t}\frac{\widetilde{U}^+_n(\Phi_n - 1)}{\Phi_n} = -U^-_{n+1}\widetilde{U}^-_n - \frac{1}{t}(n+k+1)\widetilde{U}^-_{n+1}U^-_n. \tag{4.3}$$

Now solve (3.18) for $U^-_{n+1}\widetilde{U}^-_n$ and insert the result in the right hand side of (4.3). The result can be written as

$$k\widetilde{U}^+_{n+1} - (2n+2+k)U^-_n\widetilde{U}^-_{n+1} = t\,\frac{\Phi'_n}{\Phi_n} - n\frac{\Phi_n - 1}{\Phi_n} - \frac{\widetilde{U}^+_n(\Phi_n - 1)}{\Phi_n} + t - \widetilde{U}^+_n = \mathcal{E}, \tag{4.4}$$

say. Noting that

$$U^-_n\,\widetilde{U}^-_{n+1} = \widetilde{U}^+_n - \widetilde{U}^+_{n+1} \tag{4.5}$$

we have

$$\left(U^-_n\,\widetilde{U}^-_{n+1}\right)' = \Phi_n - \Phi_{n+1}$$

by (3.25). Therefore differentiating the left side of (4.4) gives

$$\mathcal{E}' = -(2n+2+k)\,\Phi_n + 2(k+n+1)\,\Phi_{n+1},$$

$$2(k+n+1)\,\Phi_{n+1} = \mathcal{E}' + (2n+2+k)\,\Phi_n.$$

Computing $\mathcal{E}'$ using the right side of (4.4) we obtain the representation

$$2(n+k+1)\,\Phi_{n+1} = 2 + 2(n+k)\,\Phi_n - (n + \widetilde{U}^+_n - \Phi_n)\frac{\Phi'_n}{\Phi_n^2} - t\,\frac{(\Phi'_n)^2}{\Phi_n^2} + t\,\frac{\Phi''_n}{\Phi_n}. \tag{4.6}$$

Simply integrating the preceding equation using (3.25) gives

$$2(n+k+1)\,\widetilde{U}^+_{n+1} = t\,\frac{\Phi'_n}{\Phi_n} - n\,\frac{\Phi_n - 1}{\Phi_n} + 2(k+n)\,\widetilde{U}^+_n + \frac{\widetilde{U}^+_n}{\Phi_n} + t. \tag{4.7}$$

We now write out the last relation from which the differential equation will follow.

$$\Phi_{n+1} = 1 - U_{n+1}^{-} \widetilde{U}_{n+1}^{-} = 1 - \frac{(U_{n+1}^{-} \widetilde{U}_{n}^{-})\,(U_{n}^{-} \widetilde{U}_{n+1}^{-})}{U_{n}^{-} \widetilde{U}_{n}^{-}} = 1 - \frac{(U_{n+1}^{-} \widetilde{U}_{n}^{-})\, U_{n}^{-} \widetilde{U}_{n+1}^{-})}{1 - \Phi_n}. \tag{4.8}$$

Use (4.5) and (4.7) to express $U_n^-\, \widetilde{U}_{n+1}^-$ in terms of $\widetilde{U}_n^+$, Φ_n and Φ_n'. Similarly use (4.3) (4.5) and (4.7) to express $\widetilde{U}_n^-\, U_{n+1}^-$ in terms of the same quantities. Finally, on the left hand side of (4.8) use (4.6). The result is a third-order differential equation for $\widetilde{U}_n^+$. (Recall (3.25).) This third-order equation is

$$w''' = \frac{1}{2}\left(\frac{1}{w'} + \frac{1}{w'-1}\right)(w'')^2 - \frac{1}{t}\, w'' + \frac{2(k+n)}{t}\, w'$$
$$- \frac{2(k+n)}{t}\,(w')^2 + \frac{t+n}{2t^2}\,(n - t + 2\,w) - \frac{(n+w)^2}{2t^2 w'} - \frac{(t-w)^2}{2t^2(w'-1)}. \tag{4.9}$$

Cosgrove tells us[7] that the equation integrates to

$$t^2\, (w'')^2 = -4(k+n)\, t\, (w')^3 + \left\{4(k+n)\, w + t^2 + 2(2k+3n)\, t + n^2\right\} (w')^2$$
$$- \left\{2(t + 2k + 3n)\, w + 2n\, t + 2n^2\right\}\, w' + (w+n)^2. \tag{4.10}$$

The $\sigma = \sigma(t)$ form of Painlevé V as given by equation (C.45) in Jimbo-Miwa [19] (see also [27]) is, after changing σ to $-\sigma$ and taking the special parameter values $\nu_0 = \nu_1 = 0$, $\nu_2 = k,\ \nu_3 = k+n$,

$$(t\,\sigma'')^2 = \left\{\sigma - t\,\sigma' - 2\,(\sigma')^2 + (2k+n)\,\sigma'\right\}^2 - 4\,(\sigma')^2\,(\sigma' - k)\,(\sigma' - k - n). \tag{4.11}$$

If

$$w = t - \frac{\sigma}{(k+n)},$$

then (4.10) and (4.11) are equivalent. Notice that since $w = \widetilde{U}_n^+$, (3.17) says that $\sigma = k\,t - t\, U_n^+$ and therefore by (3.22)

$$\sigma = -t\, \frac{d}{dt} \log\left(e^{-k\,t}\, D_n(t)\right)$$

and therefore

$$e^{-k\,t}\, D_n(t) = \exp\left(-\int_0^t \frac{\sigma(t')}{t'}\, dt'\right). \tag{4.12}$$

This, with Theorem 1, gives Theorem 2 for G_I. For G_D it is simply a matter of changing k to $-k$ and t to $-t$.

[7] Cosgrove, in his analysis of certain third-order differential equations, has shown that the third order differential equation of Chazy Class I (see (A.3) in [10]) can be integrated to a second order and second degree "master Painlevé equation" (see (A.21) in [10]). This master Painlevé equation, called SD-I in [9], contains all the Painlevé equations I–VI. Our (4.9) is a special case of Cosgrove's (A.3). Carrying out this reduction [11] in this special case results in (4.10). Cosgrove's integration constant equals $-n^2/4$ in our case. This follows from the boundary conditions derived below in (4.14).

2. Laguerre ensemble interpretation of $D_n(t)$

In order to specify which solution of (4.11) our σ is, we must determine the boundary condition σ satisfies at $t=0$. We have by (3.25) and (3.6)

$$\frac{d\sigma}{dt} = (k+n)\, U_n^-\, \widetilde{U}_n^- = (k+n)\, \frac{V_{n+1}^-\, \widetilde{V}_{n+1}^-}{(V_{n+1}^+)^2}.$$

Now V_{n+1}^+ is the upper-left entry of $T_{n+1}(f)^{-1}$ where, recall, $f(z)=e^{t/z}\,(1+z)^k$. As $t\to 0$ this approaches the upper-left entry of $(I+\Lambda')^{-k}$, which is clearly equal to 1. Equally clearly, the lower-left entry of the inverse has limit $\binom{-k}{n}$, so that

$$\lim_{t\to 0} V_{n+1}^- = \binom{-k}{n}.$$

To determine the behavior of $\widetilde{V}_{n+1}^-$, the the upper-right entry of $T_{n+1}(f)^{-1}$, we write

$$T_{n+1}(f) = (I+\Lambda')^k + \sum_{p>0,\, q\geq 0} \frac{t^p}{p!}\binom{k}{q}\Lambda^{(p-q)}, \tag{4.13}$$

where $\Lambda^{(j)}$ denotes Λ^j if $j\leq 0$ and $(\Lambda')^{-j}$ if $j<0$. Factoring out $(I+\Lambda')^k$ and taking the inverse gives

$$T_{n+1}(f)^{-1} = \left(I + (I+\Lambda')^{-k} \sum_{p>0,\, q\geq 0} \frac{t^p}{p!}\binom{k}{q}\Lambda^{(p-q)}\right)^{-1} (I+\Lambda')^{-k}.$$

If we expand out the inverses we get a sum of products. Each product has factors of the form $t^{p_i}\Lambda^{(p_i-q_i)}$ and other factors which are nonnegative powers of Λ'. Such a product will have a nonzero upper-right entry only if $\sum(p_i+q_i)\geq n$. Therefore, since each $p_i\geq 1$, the lowest power of t which can occur is n. Moreover this power occurs only when all $q_i=0$ and all the nonegative powers of Λ' which occur in the product are 0. This means that we get the same lowest power of t in the upper-right entry of the inverse if in (4.13) we replace $(I+\Lambda')^k$ by I and in the sum we only take the terms with $q=0$. This amounts to replacing $T_{n+1}(f)$ by

$$\sum_{p\geq 0} \frac{t^p}{p!}\,\Lambda^p = e^{t\,\Lambda}.$$

The inverse of this operator is $e^{-t\,\Lambda}$ and the upper-right corner of this matrix is exactly $(-1)^n\, t^n/n!$. Thus

$$\widetilde{V}_{n+1}^- = \frac{(-t)^n}{n!} + O(t^{n+1}),$$

as $t\to 0$, and so

$$\frac{d\sigma}{dt} = (k+n)\binom{-k}{n}\frac{(-t)^n}{n!} + O(t^{n+1}) = \frac{k}{n!}\binom{n+k}{n}\, t^n + O(t^{n+1}).$$

Since $\sigma(0) = 0$,

$$\sigma(t) = \frac{k}{(n+1)!} \binom{n+k}{n} t^{n+1} + O(t^{n+2}). \tag{4.14}$$

Here is the remarkable fact: the same function σ which satisfies the equation (4.11) together with the boundary condition (4.14) gives a representation for the Fredholm determinat which equals the distribution function for the smallest eigenvalue in the Laguerre ensemble of $k \times k$ matrices associated with the weight function $x^n\, e^{-x}$. Precisely, we have

$$\text{Prob}\,(\lambda_{\min} \ge t) = \det\,(I - K_L)\,,$$

where K_L is the integral operator on $(0,\, t)$ with kernel

$$K_L(x,y) = [k(k+n)]^{1/2}\; \frac{\varphi_{L,k}(x)\,\varphi_{L,k-1}(y) - \varphi_{L,k-1}(x)\,\varphi_{L,k}(y)}{x-y}.$$

Here

$$\varphi_{L,k}(x) = \sqrt{\frac{k!}{(n+k)!}}\, x^{n/2}\, e^{-x/2}\, L_k^{(n)}(x).$$

Moreover

$$\det\,(I - K_L) = \exp\left(-\int_0^t \frac{\sigma(t')}{t'}\, dt'\right). \tag{4.15}$$

(See [35], Section VB.) It follows from this and (4.12) that

$$e^{-k\,t}\, D_n(t) = \det\,(I - K_L)\,,$$

which, with Theorem 1, is the assertion of Theorem 3.

3. Limiting distribution as $N \to \infty$

The foregoing can be restated in more concrete terms as

$$\sum_{N \ge 0} F_I(n;k,\,N)\, \frac{(kt)^N}{N!} = e^{kt}\, c_{k,\,n} \int_t^\infty \cdots \int_t^\infty \prod x_j^n\; e^{-\sum x_j}\, \Delta(x)^2\, dx_1 \cdots dx_k, \tag{4.16}$$

where $\Delta(x) = \prod_{i<j}(x_i - x_j)$ and $c_{k,\,n}$ is the normalization constant defined by

$$(c_{k,\,n})^{-1} = \int_0^\infty \cdots \int_0^\infty \prod x_j^n\; e^{-\sum x_j}\, \Delta(x)^2\; dx_1 \cdots dx_k.$$

In fact ([25], formula (17.6.5))

$$(c_{k,\,n})^{-1} = 1!\,2!\,\cdots\,k!\,\prod_{j=0}^{k-1}(n+j)!. \tag{4.17}$$

If we make the variable changes $x_j \to x_j + t$ in the integral, the right side of (4.16) becomes

$$c_{k,\,n} \int_0^\infty \cdots \int_0^\infty \prod (x_j + t)^n\; e^{-\sum x_j}\, \Delta(x)^2\, dx_1 \cdots dx_k.$$

Therefore

$$F_I(n;k,\ N) = \frac{N!}{k^N}\, c_{k,n} \int_0^\infty \cdots \int_0^\infty e^{-\sum x_j}\, \Delta(x)^2\, dx\ \frac{1}{2\pi i}\int t^{-N-1} \prod (x_j+t)^n\, dt,$$

where the inner integral is taken over a contour surrounding $t=0$ and we write dx for $dx_1 \cdots dx_k$.

Set

$$N = kn - r, \quad r = [sk\sqrt{2n}]. \tag{4.18}$$

Then

$$\begin{aligned}\frac{1}{2\pi i}\int t^{-N-1} \prod (x_j+t)^n\, dt &= \frac{1}{2\pi i}\int t^r\ \exp\left\{n \sum \log\left(1+t^{-1}\, x_j\right)\right\}\ \frac{dt}{t} \\ &= \frac{1}{2\pi i}\int \exp\left\{r\log t + n\ \left(t^{-1}\sum x_j - \frac{1}{2}t^{-2}\sum x_j^2 + \cdots\right)\right\}\ \frac{dt}{t},\end{aligned}$$

as long as the integration is over a contour where $|t| > \sum x_j$. In fact we integrate over the circle $|t| = n\sum x_j/r$. The function $r\log t + n\, t^{-1}\sum x_j$ has a critical point at $t = n\sum x_j/r$ and its real part restricted to the circle has an absolute maximum there. The rest of the exponent, $n\left(-\frac{1}{2}t^{-2}\sum x_j^2 + \cdots\right)$, is uniformly bounded on the circle and equals

$$-\frac{r^2}{2n}\frac{\sum x_j^2}{(\sum x_j)^2} + o(1)$$

at the critical point. It follows that as $n \to \infty$ we have, uniformly for all x_j,

$$\begin{aligned}\frac{1}{2\pi i}\int t^{-N-1}\prod(x_j+t)^n\, dt &\sim \exp\left\{-\frac{r^2}{2n}\,\frac{\sum x_j^2}{(\sum x_j)^2}\right\}\ \frac{1}{2\pi i}\int \exp\left\{r\log t + n\ \left(\frac{\sum x_j}{t}\right)\right\}\ \frac{dt}{t} \\ &= \exp\left\{-\frac{r^2}{2n}\ \frac{\sum x_j^2}{(\sum x_j)^2}\right\}\ \frac{(n\sum x_j)^r}{\Gamma(r+1)}.\end{aligned}$$

Thus

$$F_I(n;k,\ N) \sim \frac{N!}{k^N}\, c_{k,n}\, \frac{n^r}{\Gamma(r+1)} \int\cdots\int \exp\left\{-\frac{r^2}{2n}\ \frac{\sum x_j^2}{(\sum x_j)^2}\right\}\ \left(\sum x_j\right)^r e^{-\sum x_j}\, \Delta(x)^2\, dx. \tag{4.19}$$

Define

$$\mathcal{Z} := \{(x_j) \in \mathbf{R}^k\ :\ \sum x_j = 0\}$$

and for general $(x_j) \in \mathbf{R}^k$ write

$$y = \sum x_j, \quad x_j = -x_j' + y/k,$$

so that $(x_j') \in \mathcal{Z}$. We integrate over $\mathcal{Z}$ with Lebesgue measure and over $y \in \mathbf{R}$. Since each $x_j \geq 0$ the y integration is restricted to

$$y \geq k\ \max x_j'.$$

We find that the double integral in (4.19) equals (after changing back from x' to x)

$$\frac{1}{\sqrt{k}}\int_{\mathcal{Z}}\Delta(x)^2\,dx\int_{k\,\max x_j}^{\infty}\exp\left\{-\frac{r^2}{2n}\,\frac{\sum x_j^2+y^2/k}{y^2}\right\}y^r\,e^{-y}\,dy$$

$$=\frac{e^{-r^2/2kn}}{\sqrt{k}}\int_{\mathcal{Z}}\Delta(x)^2\,dx\int_{k\,\max x_j}^{\infty}\exp\left\{-\frac{r^2}{2n}\,\frac{\sum x_j^2}{y^2}\right\}y^r\,e^{-y}\,dy$$

$$=\frac{e^{-r^2/2kn}}{\sqrt{k}}\,(2n)^{(k^2+r)/2}\int_{\mathcal{Z}}\Delta(x)^2\,dx\int_{k\,\max x_j}^{\infty}\exp\left\{-\frac{r^2}{2n}\,\frac{\sum x_j^2}{y^2}\right\}y^r\,e^{-y\sqrt{2n}}\,dy,$$

where to obtain the last we made the substitutions $x_j\to\sqrt{2n}\,x_j,\ y\to\sqrt{2n}\,y.$

The factor $y^r\,e^{-y\sqrt{2n}}$ in the inner integral achieves its maximum on $\mathbf{R}^+$ when $y=r/\sqrt{2n}=sk+o(1)$, at which point the other factor in the integral equals $e^{-\sum x_j^2}$. Hence if $\max x_j<s$ (so that sk is interior to the range of the y integration) the inner intregral is asymptotically equal to

$$e^{-\sum x_j^2}\int_0^{\infty}y^r\,e^{-y\sqrt{2n}}\,dy=e^{-\sum x_j^2}\,(2n)^{-(r+1)/2}\,\Gamma(r+1),$$

while if $\max x_j>s$ the inner integral is o of this. Moreover the inner integral is at most

$$\int_{k\,\max x_j}^{\infty}y^r\,e^{-y\sqrt{2n}}\,dy\le e^{-k\,\max x_j}\int_0^{\infty}y^r\,e^{-y\,(\sqrt{2n}-1)}\,dy\le C\,e^{-k\,\max x_j}\,(2n)^{-(r+1)/2}\,\Gamma(r+1)$$

for a constant C independent of the x_j and n. Hence application of the dominated convergence theorem shows that if we define

$$\mathcal{Z}_s=\{x\in\mathcal{Z}:\max x_j\le s\}$$

then the double integral in (4.19) is asymptotically

$$\frac{e^{-r^2/2kn}}{\sqrt{k}}\,(2n)^{(k^2-1)/2}\,\Gamma(r+1)\int_{\mathcal{Z}_s}e^{-\sum x_j^2}\,\Delta(x)^2\,dx.$$

If we recall the definition (4.18) of r and the value of $c_{k,n}$ given by (4.17) and apply Stirling's theorem we obtain

$$F_I(n;k,\,N)\sim\gamma_k\int_{\mathcal{Z}_s}e^{-\sum x_j^2}\,\Delta(x)^2\,dx,$$

where

$$\gamma_k^{-1}=1!\,2!\,\cdots\,k!\,(2\pi)^{(k-1)/2}\,2^{-(k^2-1)/2}. \tag{4.20}$$

Equivalently,

$$\lim_{N\to\infty}\mathrm{Prob}_k\left(\frac{\ell_N^I(w)-N/k}{\sqrt{2N/k}}\le s\right)=\gamma_k\int_{\mathcal{Z}_s}e^{-\sum x_j^2}\,\Delta(x)^2\,dx. \tag{4.21}$$

The right side is the conditional probability that the largest eigenvalue of a matrix from GUE is at most s, given that the matrix has trace zero, and so Theorem 4 is established.

4. Large k asymptotics

In this section we denote by $F(s,k)$ the probability in $k\times k$ GUE that $\lambda_{\max}\le s$ and by $F^0(s,k)$ the same probability in $k\times k$ traceless GUE. Thus $F^0(s,\,k)$ is given by the right side of (4.21) while

$$F(s,\,k)=c_k\int_{-\infty}^{s}\cdots\int_{-\infty}^{s}e^{-\sum x_j^2}\,\Delta(x)^2\,dx_1\cdots dx_k,$$

where (see (4.20) and formula (17.6.7) of [25]) $c_k=\gamma_k/\sqrt{\pi}$. Using now the variable change

$$y=\sum x_j,\quad x_j=x_j'+y/k$$

and then changing x' back to x as before we find that

$$F(s,\,k)=\frac{\gamma_k}{\sqrt{\pi k}}\int_{-\infty}^{\infty}dy\int_{\mathcal{Z}_{s-y/k}}e^{-(\sum x_j^2+\frac{y^2}{k})}\,\Delta(x)^2\,dx$$

$$=\gamma_k\sqrt{\frac{k}{\pi}}\int_{-\infty}^{\infty}e^{-k\,y^2}\,dy\int_{\mathcal{Z}_{s-y}}e^{-\sum x_j^2}\,\Delta(x)^2\,dx=\sqrt{\frac{k}{\pi}}\int_{-\infty}^{\infty}e^{-k\,y^2}\,F^0(s-y,\,k)\,dy.$$

Replacing s by $\sqrt{2k}+s/\sqrt{2}k^{1/6}$ and making the variable change $y\to y/\sqrt{2}k^{1/6}$ we obtain

$$F(\sqrt{2k}+s/\sqrt{2}k^{1/6},\,k)=\frac{k^{1/3}}{\sqrt{2\pi}}\int_{-\infty}^{\infty}e^{-k^{2/3}\,y^2/2}\,F^0(\sqrt{2k}+(s-y)/\sqrt{2}k^{1/6},\,k)\,dy.$$

The factor $F^0(\sqrt{2k}+(s-y)/\sqrt{2}k^{1/6},\,k)$ in the integrand is a bounded and nonincreasing function of y whereas the integral over $|y|>\delta$ is $o(1)$ for any δ and the right side without this factor equals one. Hence for any $\delta>0$

$$F^0(\sqrt{2k}+(s-\delta)/\sqrt{2}k^{1/6},\,k)+o(1)\le F(\sqrt{2k}+s/\sqrt{2}k^{1/6},\,k)$$
$$\le F^0(\sqrt{2k}+(s+\delta)/\sqrt{2}k^{1/6},\,k)+o(1).$$

It follows from this, (1.3) and the fact that $F_2(s)$ is continuous that

$$\lim_{k\to\infty}F^0(\sqrt{2k}+s/\sqrt{2}k^{1/6},\,k)=F_2(s).\tag{4.22}$$

(A result which includes this, where $N\to\infty$ and $k\to\infty$ with $N\gg k$, can be found in [22].)

Another way of expressing (4.22) is as follows. Let ℓ_k equal the weak limit of $(\ell_N^I-N/k)/\sqrt{N}$ as $N\to\infty$. The distribution function of ℓ_k is

$$\lim_{N\to\infty}\mathrm{Prob}_k\left(\frac{\ell_N^I(w)-N/k}{\sqrt{N}}\le s\right)=F^0(s\sqrt{k/2},\,k).$$

Then (4.22) is equivalent to the statement

$$\lim_{k\to\infty}\mathrm{Prob}\left((\ell_k-2)k^{2/3}\le s\right)=F_2(s).$$

Acknowlegements

The authors have benefited from conversations and correspondence with D. Aldous, Y. Chen, P. Clarkson, C. Cosgrove, P. Diaconis, I. Gessel, C. Grinstead, A. Its, K. Johansson, J. L. Snell and R. Stanley. It is a pleasure to acknowledge this and, as well, the support provided by the National Science Foundation through grants DMS–9802122 and DMS–9732687. Finally, we wish to thank D. Eisenbud and H. Rossi for their support during the MSRI semester Random Matrix Models and Their Applications.

References

[1] D. Aldous and P. Diaconis, Longest increasing subsequences: From patience sorting to the Baik-Deift-Johansson theorem, preprint.

[2] J. Baik, P. Deift, and K. Johansson, On the distribution of the length of the longest increasing subsequence of random permutations, preprint (math.CO/9810105).

[3] J. Baik, P. Deift, and K. Johansson, On the distribution of the length of the second row of a Young diagram under Plancherel measure, preprint (math.CO/9901118).

[4] J. Baik and E. M. Rains, The asymptotics of monotone subsequences of involutions, preprint (math.CO/9905084).

[5] J. Baik and E. M. Rains, Algebraic aspects of increasing subsequences, preprint (math.CO/9905083).

[6] A. Borodin, Longest increasing subsequences of random colored permutations, preprint (math.CO/9902001).

[7] A. Borodin, A. Okounkov and G. Olshanski, Asymptotics of Plancherel measures for symmetric groups, preprint (math.CO/9905032).

[8] P. A. Clarkson and J. B. McLeod, A connection formula for the second Painlevé transcendent, *Arch. Rat. Mech. Anal.* **103** (1988), 97–138.

[9] C. M. Cosgrove and G. Scoufis, Painlevé classification of a class of differential equations of the second order and second degree, *Stud. Appl. Math.* **88** (1993), 25–87.

[10] C. M. Cosgrove, Chazy classes IX–XI of third-order differential equations, *Stud. Appl. Math.,* to appear.

[11] C. M. Cosgrove, letter to authors, March 19, 1999.

[12] P. A. Deift and X. Zhou, A steepest descent method for oscillatory Riemann-Hilbert problems: Asymptotics for the MKdV equation, *Ann. Math.* **137** (1993), 295–368.

[13] P. A. Deift and X. Zhou, Asymptotics for the Painlevé II equation, *Commun. Pure Appl. Math.* **48** (1995), 277–337.

[14] P. J. Forrester, Exact results and universal asymptotics in the Laguerre random matrix ensemble, J. Math. Phys. **35** (1994), 2539–2551.

[15] W. Fulton, *Young Tableaux: With Applications to Representation Theory and Geometry*, Cambridge University Press, 1997.

[16] I. M. Gessel, Symmetric functions and P-recursiveness, *J. Comb. Theory, Ser. A,* **53** (1990), 257–285.

[17] S. P. Hastings and J. B. McLeod, A boundary value problem associated with the second Painlevé transcendent and the Korteweg de Vries equation, *Arch. Rat. Mech. Anal.* **73** (1980), 31–51.

[18] M. Hisakado, Unitary matrix models and Painlevé III, *Mod. Phys. Letts.,* **A11** (1996), 3001–3010.

[19] M. Jimbo and T. Miwa, Monodromy preserving deformation of linear ordinary differential equations with rational coefficients. II, *Physica* **2D** (1981), 407–448.

[20] K. Johansson, The longest increasing subsequence in a random permutation and a unitary random matrix model, *Math. Research Lett.* **5** (1998), 63–82.

[21] K. Johansson, Shape fluctuations and random matrices, preprint (math.CO/9903134).

[22] K. Johansson, Discrete orthogonal polynomial ensembles and the Plancherel measure, preprint (math.CO/9906120).

[23] D. E. Knuth, Permutations, matrices and generalized Young tableaux, *Pacific J. Math.* **34** (1970), 709–727.

[24] D. E. Knuth, *The Art of Computer Programming*, Vol. 3, *Sorting and Searching*, 2nd ed., Addison-Wesley, Reading, 1998.

[25] M. L. Mehta, *Random Matrices*, 2nd ed., Academic Press, San Diego, 1991.

[26] A. M. Odlyzko, B. Poonen, H. Widom and H. S. Wilf, On the distribution of longest increasing subsequences in random permutations, unpublished notes.

[27] K. Okamoto, Studies on the Painlevé equations II. Fifth Painlevé equation P_V, *Japan. J. Math.* **13** (1987), 47–76.

[28] A. Okounkov, Random matrices and random permutations, preprint (math.CO/99-03176).

[29] E. M. Rains, Increasing subsequences and the classical groups, *Elect. J. of Combinatorics*, **5** (1998), #R12.

[30] A. Regev, Asymptotic values for degrees associated with strips of Young diagrams, *Adv. in Math.* **41** (1981), 115–136.

[31] G. de B. Robinson, On the representations of S_n, *Amer. J. Math.* **60** (1938), 745–760.

[32] C. E. Schensted, Longest increasing and decreasing subsequences, *Canad. J. Math.* **13** (1961), 179–191.

[33] R. P. Stanley, *Enumerative Combinatorics*, Vol. 2, Cambridge University Press, Cambridge, 1999.

[34] C. A. Tracy and H. Widom, Level-spacing distributions and the Airy kernel, *Commun. Math. Phys.* **159** (1994), 151–174.

[35] C. A. Tracy and H. Widom, Fredholm determinants, differential equations and matrix models, *Commun. Math. Phys.* **163** (1994), 33–72.

[36] C. A. Tracy and H. Widom, On orthogonal and symplectic ensembles, *Commun. Math. Phys.* **177** (1996), 727–754.

[37] C. A. Tracy and H. Widom, Random unitary matrices, permutations and Painlevé, preprint (math.CO/9811154).

[38] C. A. Tracy and H. Widom, Correlation functions, cluster functions and spacing distributions for random matrices, *J. Stat. Phys.* **92** (1998), 809–835.